Title

ivtobit — Tobit model with continuous endogenous covariates

Description

`ivtobit` fits tobit models where one or more of the covariates are endogenously determined. By default, `ivtobit` uses maximum likelihood estimation, but Newey's (1987) minimum chi-squared (two-step) estimator can be requested. Both estimators assume that the endogenous covariates are continuous and so are not appropriate for use with discrete endogenous covariates.

Quick start

Tobit regression of `y1` on `x` and endogenous regressor `y2` that is instrumented by `z` where `y1` is left-censored at its observed minimum

```
ivtobit y1 x (y2 = z), ll
```

As above, but specify that `y1` is left-censored at 0 and right-censored at 20

```
ivtobit y1 x (y2 = z), ll(0) ul(20)
```

Use Newey's two-step estimator

```
ivtobit y1 x (y2 = z), ll(0) ul(20) twostep
```

As above, and show first-stage regression results

```
ivtobit y1 x (y2 = z), ll(0) ul(20) twostep first
```

Menu

Statistics > Endogenous covariates > Tobit model with endogenous covariates

Syntax

Maximum likelihood estimator

`ivtobit` *depvar* [*varlist*$_1$] `(`*varlist*$_2$ `=` *varlist*$_{iv}$`)` [*if*] [*in*] [*weight*] `,`

`ll`[`(`*#*`)`] `ul`[`(`*#*`)`] [*mle_options*]

Two-step estimator

`ivtobit` *depvar* [*varlist*$_1$] `(`*varlist*$_2$ `=` *varlist*$_{iv}$`)` [*if*] [*in*] [*weight*] `,` `twostep`

`ll`[`(`*#*`)`] `ul`[`(`*#*`)`] [*tse_options*]

varlist$_1$ is the list of exogenous variables.

varlist$_2$ is the list of endogenous variables.

varlist$_{iv}$ is the list of exogenous variables used with *varlist*$_1$ as instruments for *varlist*$_2$.

mle_options	Description
Model	
`ll`[`(`#*`)`]	left-censoring limit
`ul`[`(`#*`)`]	right-censoring limit
`mle`	use conditional maximum-likelihood estimator; the default
`constraints(`*constraints*`)`	apply specified linear constraints
SE/Robust	
`vce(`*vcetype*`)`	*vcetype* may be `oim`, `robust`, `cluster` *clustvar*, `opg`, `bootstrap`, or `jackknife`
Reporting	
`level(`*#*`)`	set confidence level; default is `level(95)`
`first`	report first-stage regression
`nocnsreport`	do not display constraints
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process
`coeflegend`	display legend instead of statistics

You must specify at least one of `ll`[`(`#*`)`] and `ul`[`(`*#*`)`].

tse_options	Description
Model	
* twostep	use Newey's two-step estimator; the default is mle
* ll[(#)]	left-censoring limit
* ul[(#)]	right-censoring limit
SE	
vce(*vcetype*)	*vcetype* may be twostep, bootstrap, or jackknife
Reporting	
level(#)	set confidence level; default is level(95)
first	report first-stage regression
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
coeflegend	display legend instead of statistics

*twostep is required. You must specify at least one of ll[(#)] and ul[(#)].

varlist$_1$ and *varlist*$_{\mathrm{iv}}$ may contain factor variables; see [U] **11.4.3 Factor variables**.

depvar, *varlist*$_1$, *varlist*$_2$, and *varlist*$_{\mathrm{iv}}$ may contain time-series operators; see [U] **11.4.4 Time-series varlists**.

bootstrap, by, jackknife, rolling, statsby, and svy are allowed; see [U] **11.1.10 Prefix commands**. fp is allowed with the maximum likelihood estimator.

Weights are not allowed with the bootstrap prefix; see [R] **bootstrap**.

vce(), first, twostep, and weights are not allowed with the svy prefix; see [SVY] **svy**.

fweights, iweights, and pweights are allowed with the maximum likelihood estimator. fweights are allowed with Newey's two-step estimator. See [U] **11.1.6 weight**.

coeflegend does not appear in the dialog box.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options for ML estimator

Model

ll[(#)] and ul[(#)] indicate the lower and upper limits for censoring, respectively. You may specify one or both. Observations with *depvar* $\leq$ ll() are left-censored; observations with *depvar* $\geq$ ul() are right-censored; and remaining observations are not censored. You do not have to specify the censoring values at all. It is enough to type ll, ul, or both. When you do not specify a censoring value, ivtobit assumes that the lower limit is the minimum observed in the data (if ll is specified) and that the upper limit is the maximum (if ul is specified).

mle requests that the conditional maximum-likelihood estimator be used. This is the default.

constraints(*constraints*); see [R] **estimation options**.

SE/Robust

vce(*vcetype*) specifies the type of standard error reported, which includes types that are derived from asymptotic theory (oim, opg), that are robust to some kinds of misspecification (robust), that allow for intragroup correlation (cluster *clustvar*), and that use bootstrap or jackknife methods (bootstrap, jackknife); see [R] ***vce_option***.

Reporting

level(#); see [R] **estimation options**.

first requests that the parameters for the reduced-form equations showing the relationships between the endogenous variables and instruments be displayed. For the two-step estimator, first shows the first-stage regressions. For the maximum likelihood estimator, these parameters are estimated jointly with the parameters of the tobit equation. The default is not to show these parameter estimates.

nocnsreport; see [R] **estimation options**.

display_options: noci, nopvalues, noomitted, vsquish, noemptycells, baselevels, allbaselevels, nofvlabel, fvwrap(#), fvwrapon(*style*), cformat(%*fmt*), pformat(%*fmt*), sformat(%*fmt*), and nolstretch; see [R] **estimation options**.

Maximization

maximize_options: difficult, technique(*algorithm_spec*), iterate(#), [no]log, trace, gradient, showstep, hessian, showtolerance, tolerance(#), ltolerance(#), nrtolerance(#), nonrtolerance, and from(*init_specs*); see [R] **maximize**.

Setting the optimization type to technique(bhhh) resets the default *vcetype* to vce(opg).

The following option is available with ivtobit but is not shown in the dialog box:

coeflegend; see [R] **estimation options**.

Options for two-step estimator

Model

twostep is required and requests that Newey's (1987) efficient two-step estimator be used to obtain the coefficient estimates.

ll[(#)] and ul[(#)] indicate the lower and upper limits for censoring, respectively. You may specify one or both. Observations with *depvar* $\leq$ ll() are left-censored; observations with *depvar* $\geq$ ul() are right-censored; and remaining observations are not censored. You do not have to specify the censoring values at all. It is enough to type ll, ul, or both. When you do not specify a censoring value, ivtobit assumes that the lower limit is the minimum observed in the data (if ll is specified) and that the upper limit is the maximum (if ul is specified).

SE

vce(*vcetype*) specifies the type of standard error reported, which includes types that are derived from asymptotic theory (twostep) and that use bootstrap or jackknife methods (bootstrap, jackknife); see [R] ***vce_option***.

Reporting

level(#); see [R] **estimation options**.

first requests that the parameters for the reduced-form equations showing the relationships between the endogenous variables and instruments be displayed. For the two-step estimator, first shows the first-stage regressions. For the maximum likelihood estimator, these parameters are estimated jointly with the parameters of the tobit equation. The default is not to show these parameter estimates.

display_options: `noci`, `nopvalues`, `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(`*style*`)`, `cformat(%`*fmt*`)`, `pformat(%`*fmt*`)`, `sformat(%`*fmt*`)`, and `nolstretch`; see [R] **estimation options**.

The following option is available with `ivtobit` but is not shown in the dialog box:

`coeflegend`; see [R] **estimation options**.

Remarks and examples

stata.com

`ivtobit` fits models with censored dependent variables and endogenous covariates. You can use it to fit a tobit model when you suspect that one or more of the covariates is correlated with the error term. `ivtobit` is to tobit what `ivregress` is to linear regression analysis; see [R] **ivregress** for more information.

Formally, the model is

$$y_{1i}^* = \boldsymbol{y}_{2i}\boldsymbol{\beta} + \boldsymbol{x}_{1i}\boldsymbol{\gamma} + u_i$$
$$\boldsymbol{y}_{2i} = \boldsymbol{x}_{1i}\boldsymbol{\Pi}_1 + \boldsymbol{x}_{2i}\boldsymbol{\Pi}_2 + \boldsymbol{v}_i$$

where $i = 1, \dots, N$; $\boldsymbol{y}_{2i}$ is a $1 \times p$ vector of endogenous variables; $\boldsymbol{x}_{1i}$ is a $1 \times k_1$ vector of exogenous variables; $\boldsymbol{x}_{2i}$ is a $1 \times k_2$ vector of additional instruments; and the equation for $\boldsymbol{y}_{2i}$ is written in reduced form. By assumption $(u_i, \boldsymbol{v}_i) \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of structural parameters, and $\boldsymbol{\Pi}_1$ and $\boldsymbol{\Pi}_2$ are matrices of reduced-form parameters. We do not observe y_{1i}^*; instead, we observe

$$y_{1i} = \begin{cases} a & y_{1i}^* < a \\ y_{1i}^* & a \le y_{1i}^* \le b \\ b & y_{1i}^* > b \end{cases}$$

The order condition for identification of the structural parameters is that $k_2 \ge p$. Presumably, $\boldsymbol{\Sigma}$ is not block diagonal between u_i and $\boldsymbol{v}_i$; otherwise, $\boldsymbol{y}_{2i}$ would not be endogenous.

❑ Technical note

This model is derived under the assumption that $(u_i, \boldsymbol{v}_i)$ is independent and identically distributed multivariate normal for all i. The `vce(cluster` *clustvar*`)` option can be used to control for a lack of independence. As with the standard tobit model without endogeneity, if u_i is heteroskedastic, point estimates will be inconsistent.

❑

▷ Example 1: Estimation and parameter interpretation

We model the number of hours per week that high school boys spend using social media (`hsocial`). The data collection process caused the observations on the number of hours spent to be censored at 12 hours. A tobit-type model is therefore reasonable for our data.

We model each boy's number of hours spent using social media as a function of whether he has a smartphone (`sphone`), whether he has a computer at home (`computer`), the year in high school in which he is enrolled (`year`), and the hours per week he spends studying (`hstudy`).

We believe that there are unobservable variables that simultaneously affect `hstudy` and `hsocial`, which is to say that `hstudy` is endogenous. Because `hstudy` is endogenous, we must model it as well. Our model for the endogenous `hstudy` always includes the exogenous covariates used to model the outcome `hsocial`. We must also include at least one covariate in the model for the endogenous `hstudy` that was not included in the model for the outcome `hsocial`.

We use `ivtobit` with the default maximum-likelihood estimator to model the endogenous variable `hstudy` as a function of the highest educational degree attained by their parents (`pedu`), the time spent watching television (`tvhours`), and the exogenous covariates used to model `hsocial`.

```
. use http://www.stata-press.com/data/r15/smedia

. ivtobit hsocial i.sphone i.computer i.year (hstudy = tvhours i.pedu), ul(12)

Fitting exogenous tobit model

Fitting full model

Iteration 0:   log likelihood = -3240.5279
Iteration 1:   log likelihood = -3186.8824
Iteration 2:   log likelihood = -3173.1147
Iteration 3:   log likelihood = -3172.8561
Iteration 4:   log likelihood =  -3172.856

Tobit model with endogenous regressors          Number of obs     =      1,324
                                                   Uncensored     =        928
Limits: lower = -inf                               Left-censored  =          0
        upper = 12                                 Right-censored =        396

                                                Wald chi2(6)      =   11610.73
Log likelihood =  -3172.856                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      hstudy |  -.9610518   .0327204   -29.37   0.000    -1.025183   -.8969211
    1.sphone |   6.041781   .0625236    96.63   0.000     5.919237    6.164325
  1.computer |    2.51903   .0629128    40.04   0.000     2.395723    2.642337
             |
        year |
           2 |   .4439009   .0802309     5.53   0.000     .2866513    .6011505
           3 |   .8574705    .080476    10.65   0.000     .6997404    1.015201
           4 |   1.478215   .0816582    18.10   0.000     1.318168    1.638262
             |
       _cons |   8.955813   .2069144    43.28   0.000     8.550268    9.361357
-------------+----------------------------------------------------------------
corr(e.hstudy,|
  e.hsocial) |   .4667975   .0340342                      .3975006       .5308
 sd(e.hsocial)|   .9709836   .0266364                      .9201559    1.024619
  sd(e.hstudy)|   .9701792   .0188547                      .9339197    1.007847
------------------------------------------------------------------------------
Instrumented:  hstudy
Instruments:   1.sphone 1.computer 2.year 3.year 4.year tvhours 2.pedu 3.pedu
------------------------------------------------------------------------------
Wald test of exogeneity (corr = 0): chi2(1) = 135.19      Prob > chi2 = 0.0000
```

The coefficients in the table tell us how much the linear prediction for the outcome changes when there is a change in a covariate. Unlike the tobit model, where the linear prediction is the expected value of the outcome as if the data had not been censored, we need to incorporate the other parameters that were estimated by `ivtobit` to obtain effects that account for endogeneity.

We recommend that you use `margins` to estimate the effect of a covariate on the mean of the outcome given the covariates. Say we want to estimate the effect of all boys having a smartphone relative to the case where no boy has a smartphone on `hsocial`.

```
. margins, dydx(sphone) predict(ystar(.,.))
Average marginal effects                          Number of obs     =      1,324
Model VCE     : OIM
Expression    : E(hsocial), predict(ystar(.,.))
dy/dx w.r.t. : 1.sphone
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.sphone	6.254632	.0562391	111.22	0.000	6.144406	6.364859

```
Note: dy/dx for factor levels is the discrete change from the base level.
```

The effect is 6.3 more hours a week of social media usage. This is different from the coefficient in the table, which has a value of 6. See [R] **ivtobit postestimation** for additional examples.

Below the table, we see a Wald test for whether the correlation between the residuals from the main equation (predicting `hstudy`) and the residuals from the auxilliary equation (predicting `hsocial`) is 0. The correlation itself is 0.47 and shown in the table as `corr(e.hstudy,e.hsocial)`. If the test statistic is not significant, there is not sufficient information in the sample to reject the null hypothesis of no endogeneity. In our example, we reject the null hypothesis that supports our choice of a tobit model that accounts for endogeneity.

◁

❑ Technical note

In the tobit model with endogenous covariates, we assume that $(u_i, \boldsymbol{v}_i)$ is multivariate normal with covariance matrix

$$\text{Var}(u_i, \boldsymbol{v}_i) = \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_u^2 & \boldsymbol{\Sigma}_{21}' \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

Using the properties of the multivariate normal distribution, $\text{Var}(u_i|\boldsymbol{v}_i) \equiv \sigma_{u|v}^2 = \sigma_u^2 - \boldsymbol{\Sigma}_{21}'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. Calculating the marginal effects on the conditional expected values of the observed and latent dependent variables and on the probability of censoring requires an estimate of σ_u^2. Unlike the default maximum-likelihood estimator, the two-step estimator identifies only $\sigma_{u|v}^2$, not σ_u^2, so only the linear prediction and its standard error are available after you have used the `twostep` option.

❑

Stored results

ivtobit, mle stores the following in e():

Scalars

e(N)	number of observations
e(N_unc)	number of uncensored observations
e(N_lc)	number of left-censored observations
e(N_rc)	number of right-censored observations
e(llopt)	minimum of *depvar* or contents of ll()
e(ulopt)	minimum of *depvar* or contents of ul()
e(k)	number of parameters
e(k_eq)	number of equations in e(b)
e(k_eq_model)	number of equations in overall model test
e(k_aux)	number of auxiliary parameters
e(k_dv)	number of dependent variables
e(df_m)	model degrees of freedom
e(ll)	log likelihood
e(N_clust)	number of clusters
e(endog_ct)	number of endogenous covariates
e(p)	model Wald p-value
e(p_exog)	exogeneity test Wald p-value
e(chi2)	model Wald χ^2
e(chi2_exog)	Wald χ^2 test of exogeneity
e(rank)	rank of e(V)
e(ic)	number of iterations
e(rc)	return code
e(converged)	1 if converged, 0 otherwise

Macros

e(cmd)	ivtobit
e(cmdline)	command as typed
e(depvar)	name of dependent variable
e(instd)	instrumented variables
e(insts)	instruments
e(wtype)	weight type
e(wexp)	weight expression
e(title)	title in estimation output
e(clustvar)	name of cluster variable
e(chi2type)	Wald; type of model χ^2 test
e(vce)	*vcetype* specified in vce()
e(vcetype)	title used to label Std. Err.
e(method)	ml
e(opt)	type of optimization
e(which)	max or min; whether optimizer is to perform maximization or minimization
e(ml_method)	type of ml method
e(user)	name of likelihood-evaluator program
e(technique)	maximization technique
e(properties)	b V
e(estat_cmd)	program used to implement estat
e(predict)	program used to implement predict
e(footnote)	program used to implement the footnote display
e(marginsok)	predictions allowed by margins
e(marginsprop)	signals to the margins command
e(asbalanced)	factor variables fvset as asbalanced
e(asobserved)	factor variables fvset as asobserved

Matrices

e(b)	coefficient vector
e(Cns)	constraints matrix
e(ilog)	iteration log (up to 20 iterations)
e(gradient)	gradient vector
e(Sigma)	$\widehat{\boldsymbol{\Sigma}}$
e(V)	variance–covariance matrix of the estimators
e(V_modelbased)	model-based variance

Functions

e(sample)	marks estimation sample

ivtobit, twostep stores the following in e():

Scalars

e(N)	number of observations
e(N_unc)	number of uncensored observations
e(N_lc)	number of left-censored observations
e(N_rc)	number of right-censored observations
e(llopt)	contents of ll()
e(ulopt)	contents of ul()
e(df_m)	model degrees of freedom
e(df_exog)	degrees of freedom for χ^2 test of exogeneity
e(p)	model Wald p-value
e(p_exog)	exogeneity test Wald p-value
e(chi2)	model Wald χ^2
e(chi2_exog)	Wald χ^2 test of exogeneity
e(rank)	rank of e(V)

Macros

e(cmd)	ivtobit
e(cmdline)	command as typed
e(depvar)	name of dependent variable
e(instd)	instrumented variables
e(insts)	instruments
e(wtype)	weight type
e(wexp)	weight expression
e(chi2type)	Wald; type of model χ^2 test
e(vce)	*vcetype* specified in vce()
e(method)	twostep
e(properties)	b V
e(estat_cmd)	program used to implement estat
e(predict)	program used to implement predict
e(footnote)	program used to implement the footnote display
e(marginsok)	predictions allowed by margins
e(marginsprop)	signals to the margins command
e(asbalanced)	factor variables fvset as asbalanced
e(asobserved)	factor variables fvset as asobserved

Matrices

e(b)	coefficient vector
e(Cns)	constraints matrix
e(V)	variance–covariance matrix of the estimators
e(V_modelbased)	model-based variance

Functions

e(sample)	marks estimation sample

Methods and formulas

The estimation procedure used by `ivtobit` is similar to that used by `ivprobit`. For compactness, we write the model as

$$y_{1i}^* = \boldsymbol{z}_i\boldsymbol{\delta} + u_i \tag{1a}$$
$$\boldsymbol{y}_{2i} = \boldsymbol{x}_i\boldsymbol{\Pi} + \boldsymbol{v}_i \tag{1b}$$

where $\boldsymbol{z}_i = (\boldsymbol{y}_{2i}, \boldsymbol{x}_{1i})$, $\boldsymbol{x}_i = (\boldsymbol{x}_{1i}, \boldsymbol{x}_{2i})$, $\boldsymbol{\delta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$, and $\boldsymbol{\Pi} = (\boldsymbol{\Pi}_1', \boldsymbol{\Pi}_2')'$. We do not observe y_{1i}^*; instead, we observe

$$y_{1i} = \begin{cases} a & y_{1i}^* < a \\ y_{1i}^* & a \le y_{1i}^* \le b \\ b & y_{1i}^* > b \end{cases}$$

$(u_i, \boldsymbol{v}_i)$ is distributed multivariate normal with mean zero and covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_u^2 & \boldsymbol{\Sigma}_{21}' \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

Using the properties of the multivariate normal distribution, we can write $u_i = \boldsymbol{v}_i'\boldsymbol{\alpha} + \epsilon_i$, where $\boldsymbol{\alpha} = \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$; $\epsilon_i \sim N(0; \sigma_{u|v}^2)$, where $\sigma_{u|v}^2 = \sigma_u^2 - \boldsymbol{\Sigma}_{21}'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$; and ϵ_i is independent of $\boldsymbol{v}_i$, $\boldsymbol{z}_i$, and $\boldsymbol{x}_i$.

The likelihood function is straightforward to derive because we can write the joint density $f(y_{1i}, \boldsymbol{y}_{2i}|\boldsymbol{x}_i)$ as $f(y_{1i}|\boldsymbol{y}_{2i}, \boldsymbol{x}_i)\, f(\boldsymbol{y}_{2i}|\boldsymbol{x}_i)$. We have that

$$\ln f(\boldsymbol{y}_{2i}|\boldsymbol{x}_i) = -\frac{1}{2}\left(p\ln 2\pi + \ln|\boldsymbol{\Sigma}_{22}| + \boldsymbol{v}_i'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{v}_i\right)$$

and

$$\ln f(y_{1i}|\boldsymbol{y}_{2i}, \boldsymbol{x}_i) = \begin{cases} \ln\left\{1 - \Phi\left(\frac{m_i - a}{\sigma_{u|v}}\right)\right\} & y_{1i} = a \\ -\frac{1}{2}\left\{\ln 2\pi + \ln\sigma_{u|v}^2 + \frac{(y_{1i} - m_i)^2}{\sigma_{u|v}^2}\right\} & a < y_{1i} < b \\ \ln\Phi\left(\frac{m_i - b}{\sigma_{u|v}}\right) & y_{1i} = b \end{cases}$$

where

$$m_i = \boldsymbol{z}_i\boldsymbol{\delta} + (\boldsymbol{y}_{2i} - \boldsymbol{x}_i\boldsymbol{\Pi})\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$$

and $\Phi(\cdot)$ is the normal distribution function so that the log likelihood for observation i is

$$\ln L_i = w_i\left\{\ln f(y_{1i}|\boldsymbol{y}_{2i}, \boldsymbol{x}_i) + \ln f(\boldsymbol{y}_{2i}|\boldsymbol{x}_i)\right\}$$

where w_i is the weight for observation i or one if no weights were specified. Instead of estimating $\sigma_{u|v}$ and σ_v directly, we estimate $\ln\sigma_{u|v}$ and $\ln\sigma_v$.

With maximum likelihood estimation, this command supports the Huber/White/sandwich estimator of the variance and its clustered version using `vce(robust)` and `vce(cluster` *clustvar*`)`, respectively. See [P] **_robust**, particularly *Maximum likelihood estimators* and *Methods and formulas*.

The maximum likelihood version of `ivtobit` also supports estimation with survey data. For details on VCEs with survey data, see [SVY] **variance estimation**.

The two-step estimates are obtained using Newey's (1987) minimum chi-squared estimator. For more details on the minimum chi-squared estimator, see [R] **ivprobit**.

Acknowledgments

The two-step estimator is based on the `tobitiv` command written by Jonah Gelbach of the University of Pennsylvania Law School and the `ivtobit` command written by Joe Harkness formerly with the Institute of Policy Studies at Johns Hopkins University.

References

Finlay, K., and L. M. Magnusson. 2009. Implementing weak-instrument robust tests for a general class of instrumental-variables models. *Stata Journal* 9: 398–421.

Miranda, A., and S. Rabe-Hesketh. 2006. Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal, and count variables. *Stata Journal* 6: 285–308.

Newey, W. K. 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics* 36: 231–250.

Also see

[R] **ivtobit postestimation** — Postestimation tools for ivtobit

[R] **gmm** — Generalized method of moments estimation

[R] **ivprobit** — Probit model with continuous endogenous covariates

[R] **ivregress** — Single-equation instrumental-variables regression

[R] **regress** — Linear regression

[R] **tobit** — Tobit regression

[ERM] **eintreg** — Extended interval regression

[SVY] **svy estimation** — Estimation commands for survey data

[XT] **xtintreg** — Random-effects interval-data regression models

[XT] **xttobit** — Random-effects tobit models

[U] 20 Estimation and postestimation commands